

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-067351

(43)Date of publication of application : 07.03.2003

(51)Int.Cl.

G06F 15/177  
G06F 9/46

(21)Application number : 2001-258225

(71)Applicant : NEC SYSTEM TECHNOLOGIES LTD

(22)Date of filing : 28.08.2001

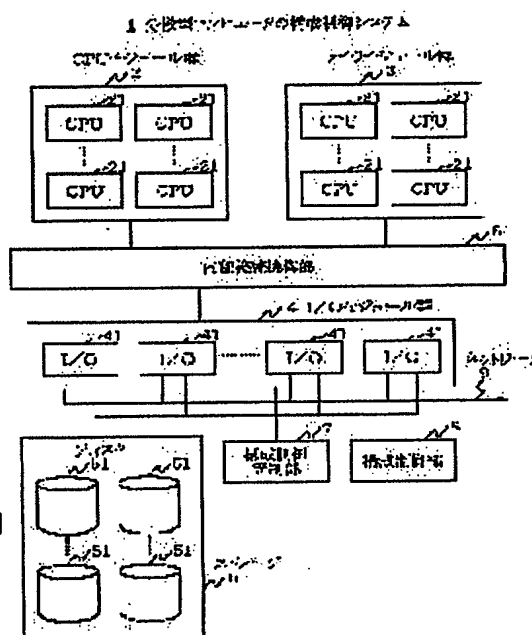
(72)Inventor : SOEJIMA YASUSHI

## (54) CONFIGURATION CONTROL SYSTEM OF DISTRIBUTED COMPUTER

(57)Abstract:

**PROBLEM TO BE SOLVED:** To perform an automatic configuration control as planned or dynamically by automatically controlling the processing capacities of partitions according to the number of modules and priority by the modularization of a distributed computer and the partitioning of selected assembly of modules.

**SOLUTION:** This configuration control system of the distributed computer comprises a CPU module group 2, a memory module group 3, an I/O module group 4, an internal connection mechanism 6 forming a communication path between the modules, a load monitoring means 61 for monitoring the loading state of the partitions formed of a network 9 connected to the I/O module and a storage 5 connected to the network 9 and having disks with a start disk, a configuration control supervising part 7 for instructing the change of the configuration of the modules of the partitions, and a configuration control part 8 for configuration-controlling the change of assembly of the modules of the partitions for the internal connection mechanism part 6.



## LEGAL STATUS

[Date of request for examination] 19.07.2002

[Date of sending the examiner's decision of rejection] 19.12.2006

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

decision of rejection]

[Date of extinction of right]

[0037] When a change in the configuration is not intentional (N in step 802), in a case of a change in which a configuration of a partition is adopted to load fluctuation (Y in step 803), if an extra module is added (Y in step 820) and the configuration is adopted to the load fluctuation (Y in step 821), the extra module is added to an applicable module in a partition in which the load fluctuation exceeded a prescribed line and change in the priority order are performed (step 824). If the configuration cannot be adopted to the load fluctuation even though the extra module is added (N in step 821), and a module is taken from other partition (Y in step 823), by taking a module from other partition, and the configuration of the applicable partition is re-formed (step 825). When a module is taken from other partition (N in step 823), the shortage remains and the existing configuration of the partition is modified (step 826). If, without adding the extra module to the partition (N in step 820), the configuration can be adopted to the load fluctuation by modification of the module (Y in step 822), the configuration and the priority order are changed within the existing partition configuration (step 828). When the configuration cannot be adopted to the load fluctuation by the modification of the module (N in step 822), the partition configuration is not changed and is left as in the existing state.

[0041] 2) Automatic configuration change by load fluctuation (a case of Fig. 7)

A case where partitions A, B, and C are present, loads of B and C fluctuates and increases, and a module is added in response to the fluctuation. When the load of the partitions B and C increase and exceed a reference operation point, load monitoring means 61, 62, and 63 respectively present in the partitions notify a configuration control management unit 7 of load states. The configuration control management unit 7, if there is a remaining extra module, instructs a configuration control unit 8 to add the extra module. The configuration control unit 8 adds the extra module to the partitions B and C as instructed. Because the configuration is adopted to the load fluctuation by the addition of the extra module the degree of priority is not changed.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-67351

(P2003-67351A)

(43) 公開日 平成15年3月7日 (2003.3.7)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード(参考)
G 0 6 F 15/177	6 7 0	G 0 6 F 15/177	6 7 0 C 5 B 0 4 5
	6 7 4		6 7 4 A 5 B 0 9 8
9/46	3 5 0	9/46	3 5 0
	3 6 0		3 6 0 C

審査請求 有 請求項の数 8 O L (全 12 頁)

(21) 出願番号 特願2001-258225(P2001-258225)

(22) 出願日 平成13年8月28日 (2001.8.28)

(71) 出願人 390001395

エヌイーシーシステムテクノロジー株式会  
社

大阪府大阪市中央区城見1丁目4番24号

(72) 発明者 副島 靖

兵庫県神戸市西区高塚台五丁目3番1号

神戸日本電気ソフトウェア株式会社内

(74) 代理人 100082935

弁理士 京本 直樹 (外2名)

Fターム(参考) 5B045 GC01

5B098 AA10 GA01 GC08 GD02 GD03

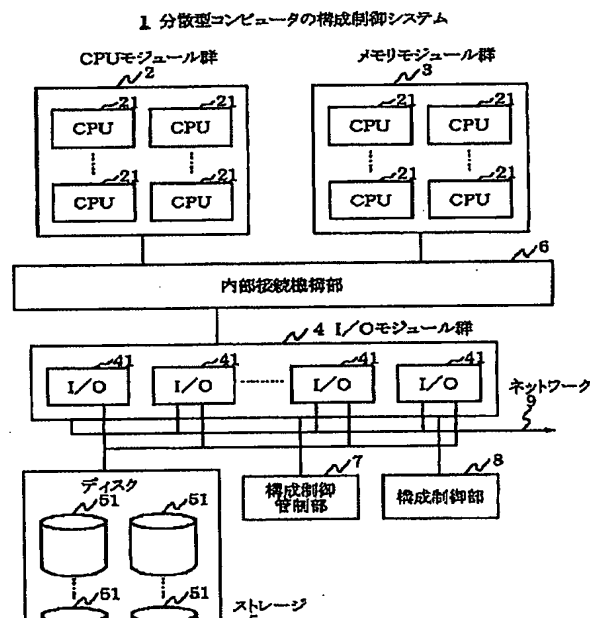
GD07 GD14

(54) 【発明の名称】 分散型コンピュータの構成制御システム

(57) 【要約】

【課題】分散型コンピュータのモジュール化と、モジュールの組み分けによるパーティション化と、によって、各パーティションの処理能力をモジュール数及び優先度によって自動調整し、計画的にあるいは動的に自動構成制御することにある。

【解決手段】CPUモジュール群2と、メモリモジュール群3と、I/Oモジュール群4と、モジュール間の通信路を形成する内部接続機構部6と、I/Oモジュールに接続するネットワーク9と、ネットワーク9に接続し、起動ディスクを含むディスクからなるストレージ5と、で構成するパーティションに対して、負荷状況を監視する負荷監視手段61と、パーティションのモジュールの構成変更指示をする構成制御管理部7と、パーティションのモジュール組み込み変更を内部接続機構部6に対して構成制御する構成制御部8と、を備える。



【0004】これまでの分散型コンピュータシステムにおける負荷制御は、各コンピュータの負荷状況を見ながら、全コンピュータの負荷に偏りが出ないようにジョブやトランザクションをコンピュータに分配することで、コンピュータの処理量を平準化して、処理能力を効率よく使い、かつ各ジョブやトランザクションの処理時間にバラツキが出ないようにするものである。このようなジョブやトランザクションを配分する負荷バランス制御は、投入しようとするジョブやトランザクションがどの程度コンピュータに負荷をかけるのかは予測できることが条件である。またジョブやトランザクションの大きさが比較的揃っていて、負荷のバラツキが小さい場合は問題にならないが、負荷に大きなバラツキがある場合はコンピュータ負荷に偏りが生じる。

【0005】この負荷の偏りを緩和するための手法として、振り分け論理にラウンドロビン形式を併用する方法があるが、効果には限界がある。また、処理途中のトランザクションやジョブをマイグレーションする手法があるが、処理途中の何時でもマイグレーションできる訳ではなく、マイグレーション可能なポイントになるまで待たされる。その間、並行処理動作しているジョブやトランザクションの処理時間は遅延してしまう。また、当該ジョブやトランザクションが必要とするコンピュータの処理能力が、各コンピュータの処理能力を超えていたなら、マイグレーションしても何ら効果がない。上述のように、従来の負荷制御は、ジョブやトランザクションを配分して、コンピュータ負荷の偏りを平準にすることであって、コンピュータ資源のハードウェアモジュールを自動的に構成制御し、処理能力を増減させて、負荷に適應させるものではなかった。また、ハードウェアを負荷対応に構成制御する場合は、構成制御部に操作コンソールを接続して人手操作を行うのが通常で、自動的に構成制御するものではなかった。

【0006】

【発明が解決しようとする課題】同種類のコンピュータ資源をモジュール化し、モジュールを組み分けてなるパーティションを自動構成制御して負荷に適應させる分散型コンピュータには、

1) コンピュータ資源を組み分けてなるパーティションの規模を調整できる。

【0007】2) 負荷変動に自動追従する構成制御によって、モジュール数を変更できる。

【0008】3) 運用計画に合せて構成制御して、パーティションの再構成やモジュール数の変更ができる。

【0009】4) 上記2)、3)の構成制御で各パーティションの優先度を考慮した制御ができる。

【0010】5) 各モジュールを均等に使用するためのローティション制御ができる。等の課題を解決すること

モジュールを組み合わせてなるコンピュータのモジュール数を増減させて、処理能力を負荷変動に適應して自動調整することにある。

【0012】更に、他の目的は、分散型コンピュータのモジュールを組み分けてパーティション化し、各パーティションのモジュール数と優先度を調整して、単独で異なる業務を処理させることにある。

【0013】更に、他の目的は、負荷変動に、動的に適應する構成制御と、運用に基づく計画的な構成変更に対処する構成制御と、を自動的に実施することにある。

【0014】

【課題を解決するための手段】そのため、この発明の、CPUモジュール群と、メモリモジュール群と、I/Oモジュール群と、前記モジュール間の通信路を形成する内部接続機構部と、前記I/Oモジュールに接続するネットワークと、前記ネットワークに接続するディスクからなるストレージと、を有してなる分散型コンピュータにあって、1つ以上のCPUモジュールと、1つ以上のメモリモジュールと、1つ以上のI/Oモジュールと、を前記内部接続機構部を介して組み込み、処理実行単位を形成するパーティションを1つ以上備え、前記パーティションは、前記I/Oモジュールを介してネットワークに接続し、所定のオペレーティングシステムを格納する起動ディスクと、前記パーティション内の各モジュールの負荷状況を監視する負荷監視手段と、を備え、前記ネットワークは、前記負荷状況の通知、あるいは運用計画に基づく前記パーティションのモジュールの構成変更指示をする構成制御管制部と、前記構成制御管制部から前記構成変更指示を受けて、前記内部接続機構部に前記パーティションのモジュール組み込みの構成変更指示をする構成制御部と、を備えてなることを特徴とする。

【0015】更に、前記内部接続機構部は、通信路を形成するスイッチ回路と、前記スイッチ回路を制御するファームウェアと、を備え、前記構成制御部の構成制御指示に基づいて、前記ファームウェアを作動させることを特徴とする。

【0016】更に、前記構成変更指示は、パーティションに組み入れる、CPUモジュールと、メモリモジュールと、I/Oモジュールの構成を定義する定義情報であることを特徴とする。

【0017】更に、前記負荷監視手段は、各パーティションに常駐して各モジュールの負荷状況を監視し、定期的に構成制御管制部に負荷状況を通知するエージェントであることを特徴とする。

【0018】更に、前記構成制御管制部は、パーティションの負荷監視手段に定期的に問い合わせる負荷状況を把握し、パーティション毎に、優先度、モジュール組み入れの基準並びに上限及び下限を設定して、パーティシ

のオペレーティングシステムをパーティションにロードし起動する。パーティション構成の定義情報は、共用メモリ上に存在し、それを各モジュールが参照して構成を認識する。

【0027】次に、この実施例の動作について説明する。まず、この発明における、構成制御の機能的動作を説明する。

#### 【0028】1) 計画的構成変更

コンピュータシステムは、負荷が時間帯によって規則的に変動する(日々の時間帯、曜日、週末処理、月末処理、期末処理、年度末処理等)。このような場合は、負荷変動に適応して構成制御するのではなく、計画的な時間割で、予め最適な構成にパーティションを構成変更する。この実施例では、時刻、日時、曜日、月末、年末、年度末、等々を契機に計画的な構成変更することができる。構成変更には、パーティションそのものの動作を継続しながらモジュール構成を増強又は縮小する場合と、一旦パーティションを崩して予備に戻し新たなパーティションを再構成して起動する場合、の2種類がある。前者の場合は負荷が存在するため、負荷変動に適応して自動構成制御する場合と同様な動作になる。計画的構成変更では、パーティションの優先度、モジュール数の上限、下限、基準、を設定することができる。なお、毎日、所定時刻に起動しているパーティションを月末、期末には臨時に増強する、というようなスケジュールによる構成制御の設定もできる。

#### 【0029】2) 負荷変動に適応するパーティションの自動構成変更

負荷に応じて自動的にモジュールの増強を行う処理である。各パーティション毎に常駐する負荷監視手段61、62、63が、CPU、メモリ、I/Oの各モジュールの負荷状況の監視を行っている。構成変更を要する程の負荷変動があった場合は、即刻、構成制御管制部7に負荷状況を伝える。構成制御管制部7自体は定期的に各パーティションの負荷状況を負荷監視手段61、62、63に問い合わせしており、絶えず負荷状況を監視している。CPU、メモリ、I/Oの各モジュールの負荷が所定のライン(作動ポイント)を越えたパーティションを見つけると、構成制御部8へ必要なモジュールの追加を指示する。

#### 【0030】3) 優先度制御

業務の優先度を「パーティションの優先度」として設定することで、処理能力を業務に適応させるモジュールの取り回しを行う。優先度の変更は、負荷変動に適応追従する場合にも適用できる。構成変更は、各パーティションの優先度と、モジュール数割り当ての上限、下限、基準の設定と、を含む。モジュールの充当は、先ず予備モジュールを割り当てる。予備がない場合は、優先度が低

ティションがない場合は、優先度の低いパーティションから下限を下回らない範囲で奪取して充当する。予備モジュールがない場合の構成制御は、次のように実施される。

【0031】(1) まず最下位優先度のパーティションからモジュールを基準数に減らしても負荷に余裕があるものを見つけ、基準分以上のモジュールを提供する。

【0032】(2) 前項(1)で全パーティションを一巡しても要求数が確保できない場合、最下位優先度のパーティションから、負荷に耐えられる下限分を残し、他を外せるモジュールとして提供する。

【0033】(3) 前項(1)、(2)を全パーティションを一巡するまで実施する。所要のモジュール数が確保できたら、各パーティションの構成をし直す。

#### 【0034】4) ローティション制御

特定のモジュールが集中的に使用されるのを避けるため、ラウンドロビン方式を基本に、使用時間を考慮したローティションを実施する。絶対ローティション回数と、モジュール毎に管理するローティション回数で制御する。モジュールには初期値として絶対ローティション回数が振られる。使用する毎にモジュールのローティション回数は1つ上がる。絶対ローティション回数は、全モジュール中の最高値のローティション回数である。最も小さい値のローティション回数のモジュールを優先的に使う。逆に予備に戻す場合は、最も値の大きいローティション回数のモジュールを優先的に戻す。ローティション回数が同じなら、使用時間の短いものを優先的に使用し、逆に予備に戻す場合には、使用時間の長いものを優先的に戻す。

【0035】上述の機能的動作を実施する構成制御管制部7の動作を説明する。図4を参照すると、パーティションの構成を変更する構成制御801は、パーティションの構成変更が計画的変更であるか(ステップ802)、負荷変動に適応するための変更であるか(ステップ803)、を判断して、それぞれの構成制御が指示される。パーティションを計画的に構成変更する場合(ステップ802のY)、先ずパーティションを組み換えるかを判定し、組み換える場合(ステップ804のY)、全パーティションの全組み換えの場合(ステップ805のY)、予備モジュールを含めて、計画のパーティションに所要のモジュールを組み込み、計画通りの構成が達成して強制的な組み換えが不必要なとき(ステップ806のN)、全パーティションの構成変更を、予備モジュールの組み込み、優先順位の変更を含めて実施する(ステップ808)。計画通りの構成が達成できず、強制的な組み換えが必要なとき(ステップ806のY)、優先順位の低位から、強制的にモジュールを奪取することを含めて、予備モジュールの組み込み、優先順位変更を実

パーティションAの負荷が上昇し基準ポイントを越えると、常駐している負荷監視手段61、62、63が構成制御管制部7に負荷状況を通知する。構成制御管制部7は、優先度が高いパーティションB、Cに、予備モジュールを追加するが、パーティションAに補充する予備モジュールが不足する。上述の優先度制御の論理で動作中の優先度が最低のパーティションDからモジュールを奪取して、パーティションAに充当する。

【0044】5) 計画的な構成変更—基準の意味—(図10の場合)

パーティションAは変えないで、パーティションB、C、DをパーティションE、F、Gに変更する場合。構成制御管制部7は優先度の高いパーティションG、F、Eの順に、モジュールの割り当てを構成制御部8に指示する。パーティションEは予備モジュールがないため、基準に未達になっているが、下限が割り当てされているため、このような割り当て結果になる。

【0045】6) 計画的な構成変更—優先度逆転—(図11の場合)

パーティションA、B、C、Dの運用を停止せずに、パーティションAを変えないで、パーティションB、C、Dの優先度、並びに上限、下限、基準の変更を行った場合。変更前は、パーティションDの負荷が上昇し動作ポイントを越え、常駐している負荷監視手段63が構成制御管制部7に負荷状況を通知したが、構成制御管制部7は予備モジュールがないため、上述の優先度制御によって動作中のパーティションA、B、Cから追加モジュールを捻出しようとするが、パーティションDの優先度が最低のため、結局何もしない。計画的な構成変更でパーティションBがパーティションDより優先度が低くなったため、上述の優先度制御の論理で動作中の優先度最低のパーティションBから、モジュールを奪取してパーティションDに充当するよう構成制御部8に指示する。

【0046】7) 計画的な構成変更—下限の扱い—(図12の場合)

パーティションAの運用を停めずに、パーティションB、C、Dを変更し、パーティションE、F、Gを構成する場合。構成制御管制部7は計画的な構成変更指示によりパーティションAの運用を停めず、パーティションB、C、Dを変更して、パーティションE、F、Gを再構成する。しかし、優先度通りモジュールを割り当てようとすると、パーティションFの基準を満たすことができないため、上述の優先度制御の論理で動作中の優先度最低のパーティションAからモジュールを奪取して、パーティションE、F、Gを構成する。

【0047】8) ローティション制御その1(図13の場合)

計画的な構成変更等を契機に、パーティション651中

【0048】構成制御管制部7は、ローティション回数が最も多いもの、ローティション回数が同じなら最も使用時間の長いモジュールを選び、予備653に戻すよう構成制御部8に指示する。CPU#2、#3、#4はローティション回数は同じだが#2、#4の方が使用時間が長いのでこの2つが予備モジュールにローティションされる。

【0049】9) ローティション制御その2(図14) 計画的な構成変更や負荷変動を契機に予備654から予備CPUモジュール2つをパーティション651に追加する時のローティション。構成制御管制部7は、ローティション回数が最も少ないモジュール、ローティション回数が同じなら最も使用時間の短いモジュールを運用に回すように構成制御部8に指示する。CPU#1はローティション回数が最も少なく、CPU#3はCPU#2、#4とローティション回数は同じだが使用時間が最も短いので、CPU#1とCPU#3をパーティションに追加する。

【0050】次に、この実施例では、パーティション自体の増強、縮小を前提としてきたが、応用としてパーティション自体を追加する変更を施すことができる。図15を参照すると、負荷バランスがよく制御されている分散型コンピュータシステムにおいて、予め予備69を設けておく。各パーティション66、67、68には、負荷監視手段を常駐させ、負荷状況を監視する。負荷が一定レベルに達したら、予備69から追加パーティション68を生成することで、システム全体の処理能力を上げる。予備69がないときは、優先度が低く負荷の低い別のパーティションから、所要のモジュールを取り外して充当する。各パーティション毎に起動ディスクが必要になるが、予め各パーティションに応じた起動ディスクを準備しておき、構成制御管制部7が追加パーティションの起動の際に、その内の1つの起動ディスク616を組み込むことで解決する。

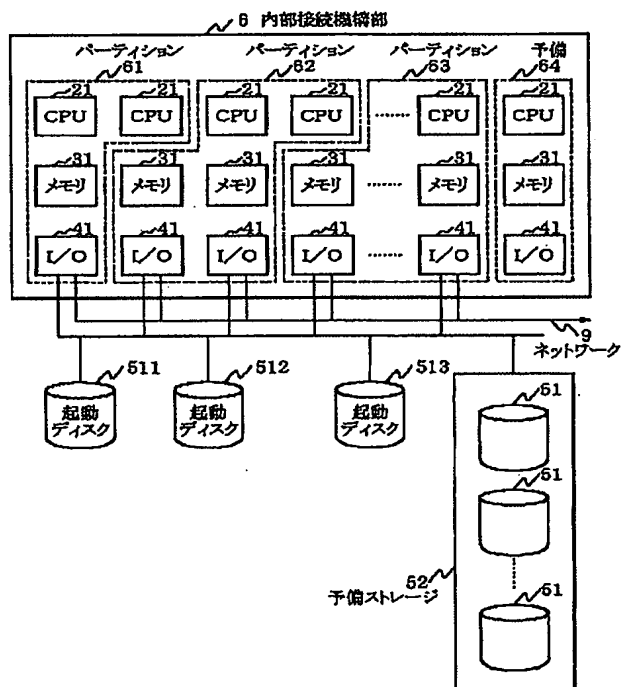
【0051】

【発明の効果】この発明の効果は、パーティションの負荷変動に適応して、自動的に構成制御することにある。その理由は、構成制御管制部が負荷監視手段と連携して負荷状況を監視して、構成制御管制部が負荷状況に見合ったパーティションのモジュール構成を決定し、構成制御部に構成変更指示をするからである。

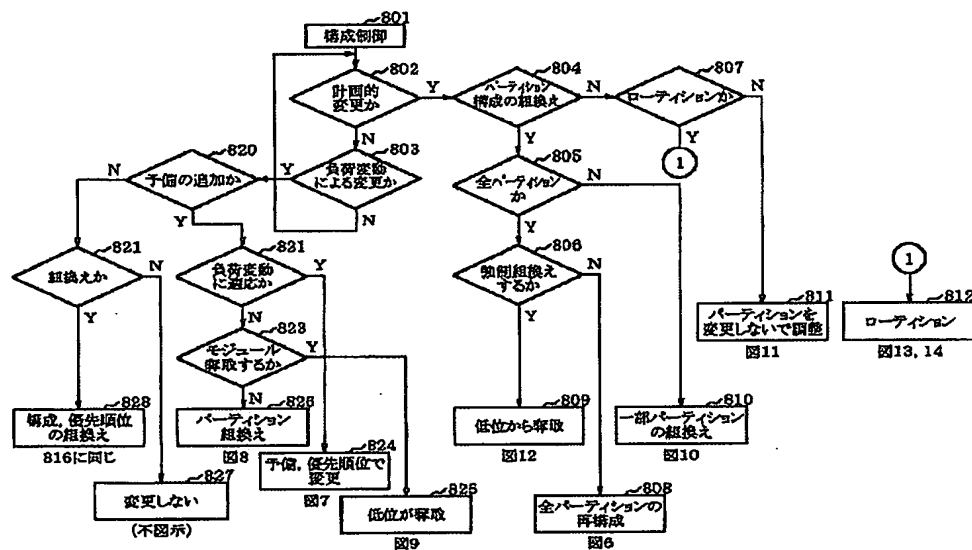
【0052】更に他の効果は、システムの運用計画に基づいて、パーティション構成を処理対象に合せて、先回りして構成変更し、運用を円滑にすることができることである。その理由は、構成制御管制部が、計画的にタイミングよくパーティションに構成変更指示できるからである。

【0053】更に他の効果は、各パーティションの優先

【圖 2】



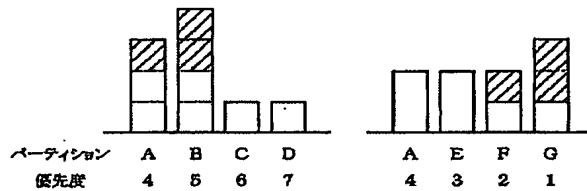
【图 4】





【図12】

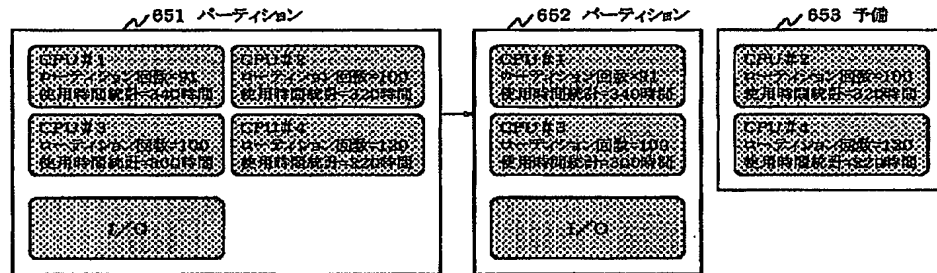
一部のパーティションの計画的再構成と  
下位のモジュール奪取



【図13】

ローテーションでパーティション縮小

パーティションのCPUを2個に縮小する場合

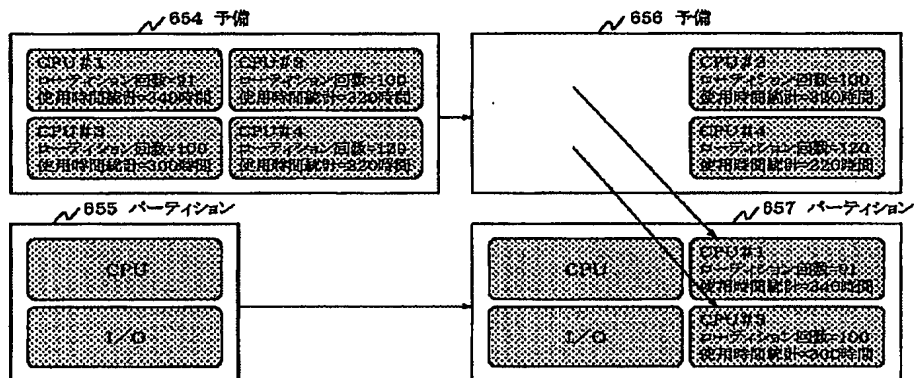


\* ローテーション回数が大きいもの、使用時間の多いものが予備に回る

【図14】

ローテーションでパーティション復元

パーティションにCPUを2個追加する場合



\* ローテーション回数が大きいもの、使用時間の多いものが予備に回る